

Monthly Report (00)

2015.12 Data Set

Monday 11th January, 2016

Prepared for

Statistics for Physical and Engineering Sciences

by

Jamie Riggs, Ph.D.

Principal Statistician
Statistics for Physical and Engineering Sciences Institute

1 Introduction

The process of reporting monthly Sunspot numbers consists of submitting individual observer's daily counts for a specific month to the AAVSO Solar Section. These data are maintained in a SQL database. The monthly data then are extracted for analysis using the R statistics package (<http://www.R-project.org/>). This report is the portion of the analysis concerned with both the raw daily average counts and the data Accuracy, Consistency, and Completeness measures for a particular month. The checks are used to scrub or filter the data to assure only error-free data are used to determine the monthly sunspot number.

This report consists of four sections: the raw daily average counts (Section 2), the known data errors (Section 3), the processed counts using a Generalized Linear Mixed Model to produce the relative sunspot numbers (Section 4), and supporting information on the model construction (Section 5).

The raw daily average of counts consist of submitted counts from all observers who provided data in the particular month. These averaged counts are reported by the day of the month, and are either from data not scrubbed or corrected data. The table captions indicate which. The errors, if any, are reported according to type.

The Error Tables section contains reported errors on missing data, inconsistencies in year and month, inconsistencies in the reported day number (1-31), seeing coding errors, number of annual observations by observer, and inconsistencies between the reported Wolf number and the calculated Wolf number from the group counts and sunspot counts, among other errors that are given in that section.

The relative sunspot numbers R_a section contains the sunspot numbers after the submitted data are scrubbed and modeled by a Generalized Linear Mixed Model (GLMM). The GLMM is a statistical model that accounts for variation due to random effects and fixed effects. For the R_a model random effects include the AAVSO observer as these observers are a selection from all possible observers, and the fixed effects include seeing conditions at one of four possible levels. More details on GLMM are available in a paper on the sunspot counts research page. The paper title is *A Generalized Linear Mixed Model for Enumerated Sunspots*.

The supporting information for the model is provided for clarification.

2 Raw Daily Average Counts

The reported raw daily average counts have been checked for errors and inconsistencies, and no known errors are present. All observers whose submissions qualify through this month's scrubbing process are represented in Figure 1 and Table 1.

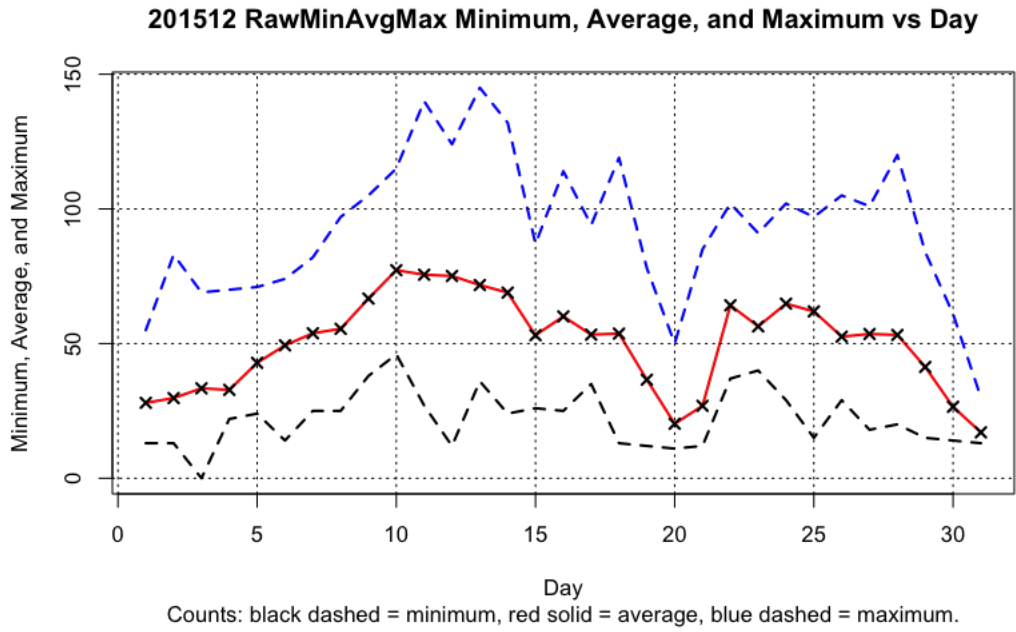


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201512 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	19.0000	13.0000	28.0000	55.0000
2.0000	26.0000	13.0000	29.7600	83.0000
3.0000	26.0000	0.0000	33.3750	69.0000
4.0000	32.0000	22.0000	32.8148	70.0000
5.0000	36.0000	24.0000	42.8387	71.0000
6.0000	34.0000	14.0000	49.3448	74.0000
7.0000	36.0000	25.0000	53.8276	82.0000
8.0000	33.0000	25.0000	55.4815	97.0000
9.0000	26.0000	38.0000	66.7083	105.0000
10.0000	29.0000	46.0000	77.2500	115.0000
11.0000	17.0000	27.0000	75.5625	140.0000
12.0000	24.0000	12.0000	75.0909	124.0000
13.0000	21.0000	36.0000	71.7368	145.0000
14.0000	27.0000	24.0000	68.9130	132.0000
15.0000	21.0000	26.0000	53.1053	87.0000
16.0000	30.0000	25.0000	60.0769	114.0000
17.0000	28.0000	35.0000	53.3200	94.0000
18.0000	22.0000	13.0000	53.7000	119.0000
19.0000	33.0000	12.0000	36.5769	78.0000
20.0000	28.0000	11.0000	20.2174	50.0000
21.0000	20.0000	12.0000	26.8947	85.0000
22.0000	22.0000	37.0000	64.2000	102.0000
23.0000	24.0000	40.0000	56.3810	91.0000
24.0000	29.0000	29.0000	64.8148	102.0000
25.0000	24.0000	15.0000	61.9500	97.0000
26.0000	28.0000	29.0000	52.5769	105.0000
27.0000	27.0000	18.0000	53.5652	101.0000
28.0000	27.0000	20.0000	53.2000	120.0000
29.0000	28.0000	15.0000	41.3200	84.0000
30.0000	30.0000	14.0000	26.5200	61.0000
31.0000	28.0000	13.0000	17.0385	30.0000

3 Error Tables

Data are for the month of December 2015. No errors were found, and hence no errors are reported.

4 Relative Sunspot Numbers

All data errors, if any, have been corrected prior to determining the following relative sunspot numbers. A Generalized Linear Mixed Model (GLMM) was constructed to provide monthly sunspot numbers (see Table 2). The GLMM treats observer as a random effect, with year, month, seeing conditions, observer rank, and dual submission to both AAVSO and SILSO as fixed effects.

Figure 2 shows the monthly R_a numbers for the years and months (ym) in Table 2. The solid cyan curve that connects the cyan X's are the GLMM model estimates given in 2. The dotted black curves on either side of the cyan curve depict a 99% confidence band about the GLMM estimates. The confidence band uses the large sample approximation based on the Gaussian distribution. The dashed red curve connecting the red O's are the SILSO values for the monthly sequence.

The tan box plots for each month are the actual observations submitted by the AAVSO observers. The heavy solid lines approximately midway in the boxes represent the count medians. The box of the box plot represents the InterQuartile Range (IQR), which depicts from the 25th through the 75th quartiles. The lower and upper whiskers extend 1.5 times the IQR below the 25th quartile, and 1.5 times the IQR above the 75th quartile. The black circles below and above the whiskers traditionally are considered outliers, but with GLMM modeling, they are observations that comprise overdispersion. Overdispersion skews the counts data from a true Poisson distribution. The GLMM adjusts for this overdispersion.

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	silso
2010.05	23.3317	22.7996	23.8638	8.4000	8.7000
2010.06	18.7725	18.2714	19.2735	11.0000	13.6000
2010.07	20.5878	20.1345	21.0411	15.2000	16.1000
2010.08	19.6390	19.1615	20.1164	18.3000	19.6000
2010.09	23.6395	23.1337	24.1454	22.8000	25.2000
2010.10	22.6863	22.1960	23.1766	21.0000	23.5000
2010.11	24.0092	23.4655	24.5529	20.9000	21.6000
2010.12	23.2504	22.5823	23.9185	13.9000	14.5000
2011.01	73.8123	72.1538	75.4708	17.7000	18.7000
2011.02	62.2569	60.8497	63.6642	29.1000	29.6000
2011.03	70.9735	69.5087	72.4383	48.0000	55.8000
2011.04	77.7501	76.0813	79.4190	47.3000	54.4000
2011.05	79.2588	77.6672	80.8504	37.3000	41.5000
2011.06	67.2470	65.8288	68.6652	35.2000	37.0000

Continued on next page

Table 2: Year Month (ym) Relative Sunspot Numbers with
 99% CI

ym	Ra	lci99	uci99	aavso	silso
2011.07	71.8854	70.3043	73.4666	41.5000	43.8000
2011.08	71.8655	70.4542	73.2769	42.4000	50.5000
2011.09	83.8470	82.7541	84.9398	73.8000	78.0000
2011.10	80.0063	78.6430	81.3695	78.9000	88.0000
2011.11	83.2675	81.5012	85.0337	84.6000	96.7000
2011.12	78.2283	76.5289	79.9277	65.8000	73.0000
2012.01	75.9138	74.4103	77.4174	55.8000	58.2000
2012.02	61.9255	60.6205	63.2304	29.2000	33.1000
2012.03	73.5452	72.2376	74.8528	53.1000	64.1000
2012.04	77.5479	75.2998	79.7959	51.4000	55.2000
2012.05	83.3710	81.9199	84.8221	61.8000	69.0000
2012.06	70.4996	69.2526	71.7466	59.7000	64.5000
2012.07	76.3396	75.0638	77.6154	64.2000	51.3000
2012.08	72.7425	71.5315	73.9535	57.7000	63.1000
2012.09	84.7088	83.2649	86.1526	57.7000	61.5000
2012.10	82.3099	80.7572	83.8626	48.3000	53.3000
2012.11	86.7557	85.0193	88.4922	56.7000	61.4000
2012.12	79.1967	77.5244	80.8690	37.4000	40.8000
2013.01	85.6092	83.9977	87.2208	63.8000	62.9000
2013.02	71.3848	69.9954	72.7743	37.8000	38.0000
2013.03	80.6145	79.1014	82.1276	50.6000	57.9000
2013.04	90.3558	88.8334	91.8781	70.6000	72.4000
2013.05	91.0572	89.4881	92.6263	77.4000	78.7000
2013.06	77.6335	76.2526	79.0143	51.0000	52.5000
2013.07	81.9350	80.6579	83.2121	57.0000	57.0000
2013.08	80.1433	78.8905	81.3961	60.0000	66.0000
2013.09	92.5046	90.9072	94.1020	34.6000	36.9000
2013.10	88.4294	86.8589	90.0000	74.5000	85.6000
2013.11	93.3190	91.3762	95.2619	73.9000	77.6000
2013.12	87.4438	85.6809	89.2066	77.8000	90.3000
2014.01	101.8252	99.6873	103.9632	77.4000	82.0000
2014.02	84.9855	83.3742	86.5968	93.9000	102.8000
2014.03	99.2206	97.5465	100.8948	80.9000	92.2000
2014.04	109.8086	107.9413	111.6759	76.9000	84.7000
2014.05	110.3197	108.5574	112.0819	72.3000	75.2000
2014.06	94.0730	92.5444	95.6016	67.2000	71.0000
2014.07	100.6935	99.0565	102.3305	72.5000	72.5000
2014.08	97.9668	96.4899	99.4437	71.2000	74.7000

Continued on next page

Table 2: Year Month (ym) Relative Sunspot Numbers with
99% CI

ym	Ra	lci99	uci99	aavso	silso
2014.09	114.0737	112.2503	115.8971	83.2000	87.6000
2014.10	108.8570	107.0336	110.6803	59.5000	60.6000
2014.11	115.4284	113.2448	117.6121	65.8000	71.1000
2014.12	105.5338	103.2673	107.8004	75.8000	78.0000
2015.01	61.8406	60.6544	63.0268	65.9000	67.0000
2015.02	51.5906	50.4024	52.7787	42.4000	44.8000
2015.03	59.1959	58.1175	60.2744	38.0000	38.4000
2015.04	66.3358	65.1665	67.5050	49.0000	54.4000
2015.05	66.3649	65.2957	67.4342	56.3000	58.8000
2015.06	56.9603	55.9979	57.9227	50.2000	68.3000
2015.07	59.6448	58.6326	60.6569	47.9000	66.4000
2015.08	59.3927	58.4244	60.3610	39.5000	64.6000
2015.09	68.9919	67.8718	70.1121	49.2000	78.1000
2015.10	65.7697	64.6474	66.8920	39.3000	61.7000
2015.11	70.2773	69.3254	71.2293	39.6000	63.2000
2015.12	63.6711	62.3729	64.9694	36.4000	57.7000

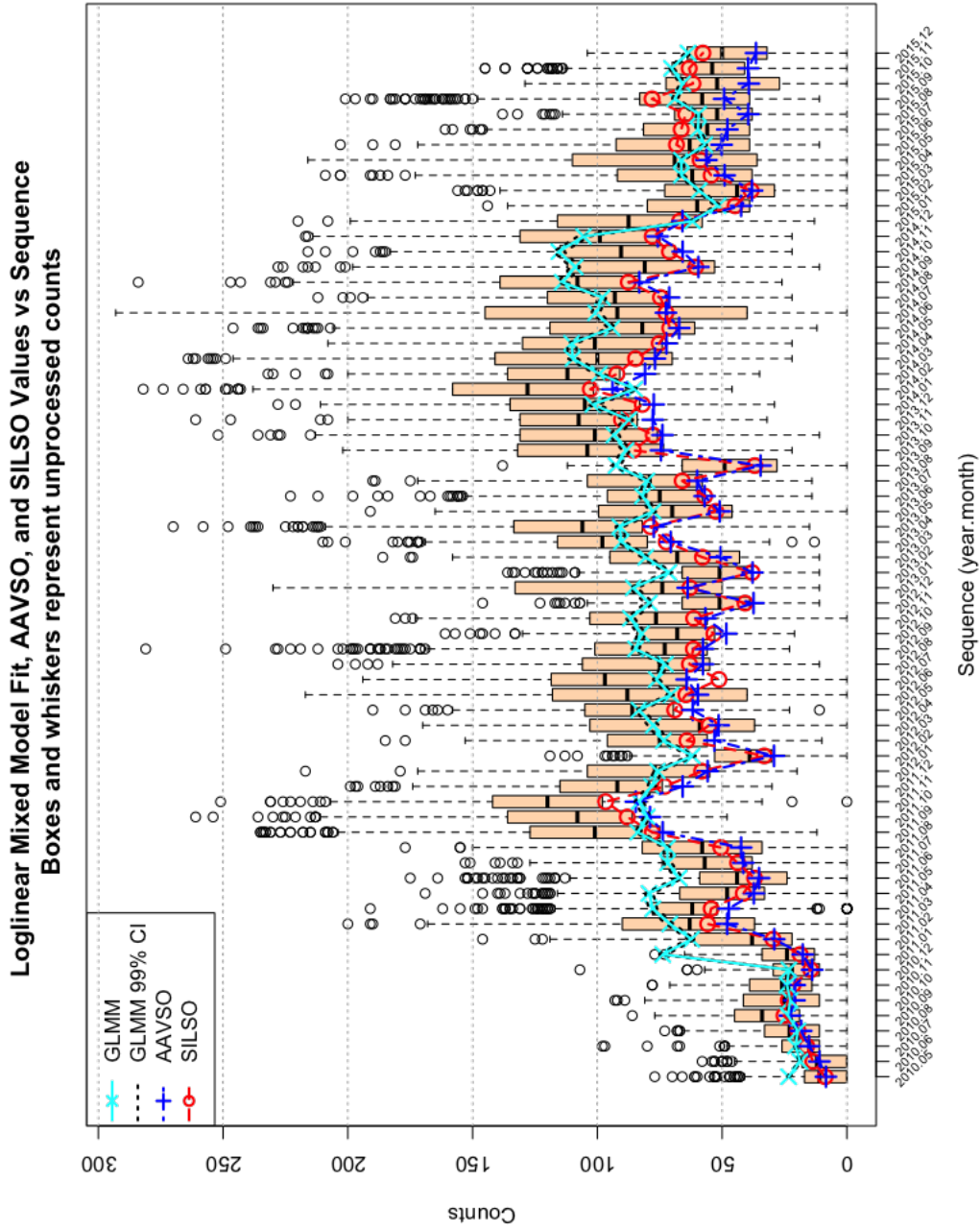


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

The GLMM parameter estimates and measures of importance in the determining the monthly R_a values are given in Table 3. The parameter estimates and levels of statistical significance are determined for the residual error size combined with the observer random effect error size. Thus, the parameter estimates are adjusted for the random effect of observer. The significance level is set at 0.05. Any $\Pr(>|z|)$ values equal to or less than 0.05 are considered statistically significant.

Table 3: 201512 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	3.1847	0.0446	71.4684	0.0000
seeF	-0.1843	0.0074	-24.7981	0.0000
seeG	-0.0974	0.0065	-15.0615	0.0000
seeP	-0.2968	0.0109	-27.1815	0.0000
r1000B	-0.0657	0.0831	-0.7911	0.4289
r1500C	0.0264	0.1270	0.2081	0.8351
r2000D	0.0743	0.1549	0.4796	0.6315
r2500E	-0.0009	0.1052	-0.0089	0.9929
r3000F	0.0610	0.1024	0.5954	0.5516
r3500G	0.1173	0.1532	0.7657	0.4438
r5000H	-0.1180	0.2120	-0.5569	0.5776
silsoy	0.1175	0.0738	1.5917	0.1115
year2011	1.2148	0.0153	79.4847	0.0000
year2012	1.2317	0.0152	80.8465	0.0000
year2013	1.3274	0.0152	87.3726	0.0000
year2014	1.5167	0.0151	100.4930	0.0000
year2015	1.0158	0.0155	65.6097	0.0000
mon2	-0.1903	0.0128	-14.9084	0.0000
mon3	-0.0532	0.0116	-4.5931	0.0000
mon4	0.0520	0.0116	4.4646	0.0000
mon5	0.0589	0.0110	5.3447	0.0000
mon6	-0.1096	0.0115	-9.4978	0.0000
mon7	-0.0473	0.0111	-4.2433	0.0000
mon8	-0.0642	0.0110	-5.8191	0.0000
mon9	0.0899	0.0106	8.4989	0.0000
mon10	0.0489	0.0112	4.3819	0.0000
mon11	0.1115	0.0113	9.8326	0.0000
mon12	0.0329	0.0120	2.7383	0.0062

The year effect levels are given as year2011, year2012, and year2013. The yearly effect is significant as $\Pr(>|z|) < 0.05$. So the year in which the observations are made is commensurate with the expected rise toward and anticipated sunspot number maximum. Similarly, the monthly effect, denoted as mon2 through mon12, is significant at the 0.05 level.

The seeing conditions account for a significant amount of deviation in sunspot numbers. The

seeing conditions are denoted as seeF (Fair), seeG (Good), and seeP (Poor), and are significant at the 0.05 level. Therefore, seeing conditions influence the reported sunspot numbers, as intuition anticipates.

The level of observer experience (denoted r1000B through r5000H, which is least to most experience) is not significant at the 0.05 significance level. It therefore does not contribute to changes in the monthly sunspot numbers.

Whether an observer contributes counts to the SILSO as well as the AAVSO (silsoy) is not significant at the 0.05 level, and hence we conclude that those observers who contribution to both institutions tend to differ from those observers contributing only to the AAVSO.

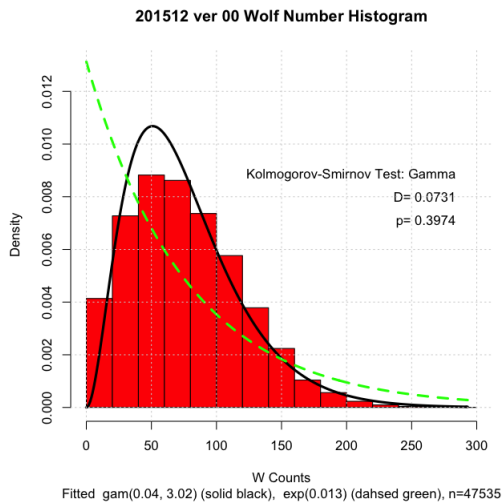
5 Supporting Information

Table 4: 201512 Summary of Sunspot Numbers

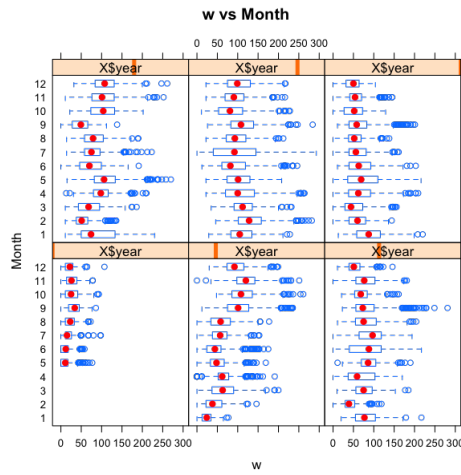
obs	jd	year	mon	day
ARAG : 2020	Min. :1721096	Min. :2010	Min. : 1.000	Min. : 1.00
CHAG : 1846	1st Qu.:2455905	1st Qu.:2011	1st Qu.: 4.000	1st Qu.: 8.00
BRAB : 1813	Median :2456408	Median :2013	Median : 7.000	Median :16.00
BROB : 1628	Mean :2456025	Mean :2013	Mean : 6.815	Mean :15.73
DUBF : 1525	3rd Qu.:2456879	3rd Qu.:2014	3rd Qu.: 9.000	3rd Qu.:23.00
HOWR : 1489	Max. :2457388	Max. :2015	Max. :12.000	Max. :31.00
(Other):37214				

Table 5: Summary of Sunspot Numbers

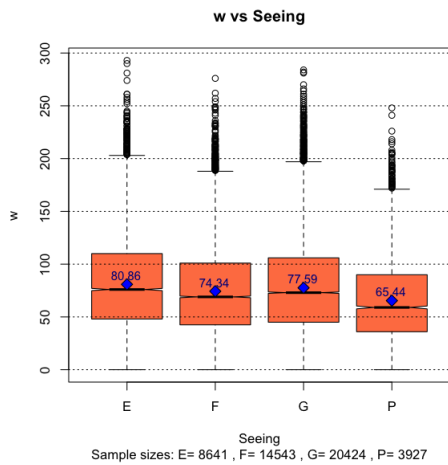
see	g	s	w	r	silso
E: 8641	Min. : 0.000	Min. : 0.00	Min. : 0.00	0000A :20409	n:32161
F:14543	1st Qu.: 3.000	1st Qu.: 12.00	1st Qu.: 44.00	3000F : 7654	y:15374
G:20424	Median : 4.000	Median : 24.00	Median : 71.00	2500E : 6150	
P: 3927	Mean : 4.666	Mean : 29.53	Mean : 76.19	3500G : 3659	
	3rd Qu.: 6.000	3rd Qu.: 42.00	3rd Qu.:104.00	1000B : 3461	
	Max. :18.000	Max. :204.00	Max. :293.00	1500C : 2834	
				(Other): 3368	



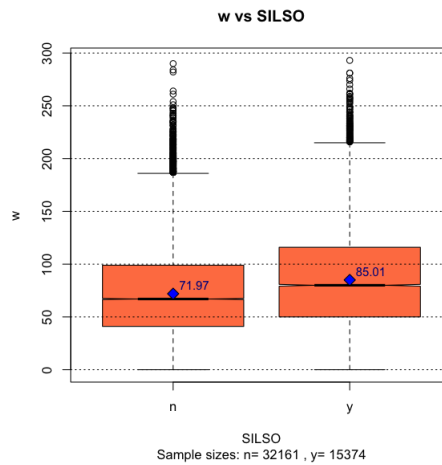
(a) Observed sunspot count histogram.



(b) Box plot of sunspot count by year and month.



(c) Box plot of sunspot count by seeing condition.



(d) Box plot of sunspot count submitted to AAVSO and SILSO.

Figure 3:

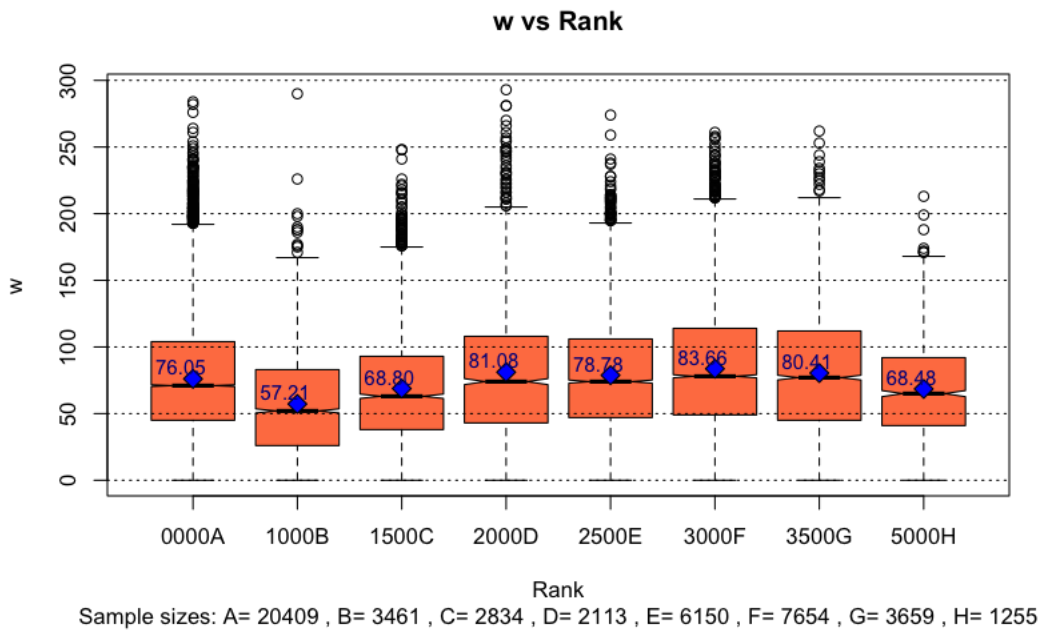


Figure 4: Box plot of sunspot count by rank.